

Tableau croisé d'effectifs

En statistique, une population est un ensemble d'individus (personnes ou objets) sur lesquels porte l'étude.

Un caractère est une propriété commune aux individus de la population.

— Sur la population des élèves d'une classe, un caractère quantitatif peut être la note obtenue au dernier devoir de français et un caractère qualitatif peut être la couleur des yeux.

Lorsqu'on étudie deux caractères sur une même population, on peut utiliser un tableau croisé (ou à double entrée) d'effectifs afin de résumer les données de l'étude.

On lit conjointement dans les lignes ou les colonnes les effectifs correspondant à chacun des caractères.

On note C_1 et C_2 , les valeurs prises par le 1^{er} caractère, D_1 et D_2 les valeurs prises par le 2^e caractère.

	C_1	C_2	Total
D_1			
D_2			
Total			

Ici, on trouve l'effectif des individus qui possèdent à la fois le caractère C_1 et le caractère D_2 .

Ici, on trouve l'effectif des individus de la population qui possèdent le caractère D_1 .

La ligne et la colonne nommées Total sont appelées les **marges** du tableau

Exemple :

Un complexe sportif étudie sa clientèle à la fin d'une journée selon l'activité pratiquée et selon le fait que le client possède un abonnement ou pas.

Les données sont résumées dans le tableau croisé ci-dessous.

	Abonné	Non abonné	Total
Activité « fitness »	75	25	100
Activité « aquatique »	90	110	200
Total	165	135	300

100 clients ont pratiqué une activité « fitness ».

135 clients ne possédaient pas d'abonnement.

90 clients ont pratiqué une activité « aquatique » et possédaient un abonnement.

Au total, il y a eu 300 clients dans ce complexe ce jour-là.

Fréquence marginale - Fréquence conditionnelle

Anna a téléchargé 20 applications sur son téléphone portable. Ces tableaux présentent la répartition de ces applications selon qu'il s'agit d'un jeu ou non, qu'elle est gratuite ou non.

Tableau des effectifs

	Jeu	Autre	Total
Gratuite	8	4	12
Payante	6	2	8
Total	14	6	20

Tableau des fréquences

	Jeu	Autre	Total
Gratuite	0,4	0,2	0,6
Payante	0,3	0,1	0,4
Total	0,7	0,3	1

Les fréquences qui se lisent dans la ligne et la colonne « Total » (en jaune) sont appelées **fréquences marginales**.

Ainsi, 0,7 est la fréquence marginale des applications de jeu dans l'ensemble des applications téléchargées par Anna.

0,6 est la fréquence marginale des applications gratuites dans l'ensemble des applications téléchargées par Anna.

On reprend la situation de l'exemple ci-dessus.

Dans cette situation, on s'intéresse uniquement aux applications gratuites. Voici les nouveaux tableaux obtenus. Dans le second tableau, les fréquences sont arrondies au centième.

Tableau des effectifs				Tableau des fréquences			
	Jeu	Autre	Total		Jeu	Autre	Total
Gratuite	8	4	12	Gratuite	0,67	0,33	1

Ainsi, parmi les applications gratuites, la fréquence des applications de jeu est 0,67. On dit que la **fréquence conditionnelle** des applications de jeu parmi les applications gratuites est 0,67.

Dans cette situation, on s'intéresse uniquement aux applications de jeu.

Voici les nouveaux tableaux obtenus. Dans le second tableau, les fréquences sont arrondies au centième.

Tableau des effectifs		Tableau des fréquences	
	Jeu		Jeu
Gratuite	8	Gratuite	0,57
Payante	6	Payante	0,43
Total	14	Total	1

Ainsi, parmi les applications de jeu, la fréquence des applications gratuites est 0,57. On dit que la fréquence conditionnelle des applications gratuites parmi les applications de jeu est 0,57.

Remarque : Dans l'exemple précédent, on peut aussi déterminer :

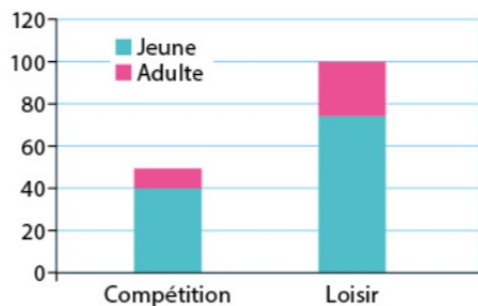
- la fréquence conditionnelle des applications de jeu parmi les applications payantes il s'agit de $\frac{6}{8} = 0,75$
- la fréquence conditionnelle des applications payantes parmi les applications de jeu ; il s'agit de $\frac{6}{14}$ soit environ 0,43.

Diagramme en barres cumulées

Voici la répartition des licenciés d'un club de natation :

	Jeune	Adulte	Total
Compétition	40	10	50
Loisir	75	25	100

Pour représenter ce tableau, on peut utiliser un diagramme en barres cumulées.



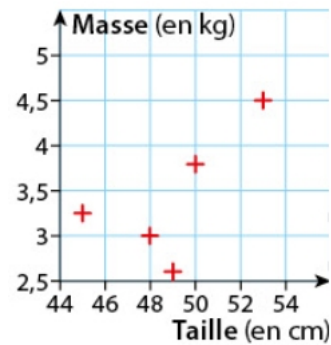
Nuage de points, point moyen

Lorsque les deux caractères étudiés sont quantitatifs, c'est-à-dire qu'ils prennent des valeurs numériques, un type de représentation graphique possible est le nuage de points : dans un repère orthogonal, l'abscisse d'un point correspond à la valeur du premier caractère étudié et son ordonnée à la valeur du second caractère.

Dans une maternité, on relève la masse et la taille de cinq nouveau-nés.

Taille t_i (en cm)	48	45	50	53	49
Masse m_i (en kg)	3	3,2	3,8	4,5	2,6

Voici ci-contre le nuage de points de coordonnées $(t_i; m_i)$ (avec i entier naturel compris entre 1 et 5) représentant cette série.



Remarque

la forme du nuage de points donne une indication d'un lien ou non entre les deux caractères étudiés. Ici, il ne semble pas y avoir de relation, affine en particulier, entre la masse et la taille de ces nouveau-nés.

Point moyen

Le point moyen G du nuage de points $(x_i; y_i)$ a pour coordonnées $(x_G; y_G)$ où x_G est la la moyenne des valeurs x_i et y_G est la la moyenne des valeurs y_i

Exemple

Chaque année, le salaire minimum (SMIC) horaire brut est revalorisé. Voici les valeurs, en euros, au 1^{er} janvier de chaque année, de 2023 à 2026.

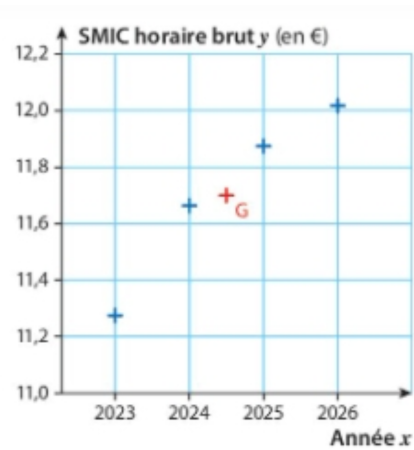
Année x	2023	2024	2025	2026
Valeur y (en €)	11,27	11,65	11,88	12,02

Ces données sont représentées par le nuage de points ci-contre. On calcule les coordonnées du point moyen G :

$$x_G = \frac{2023 + 2024 + 2025 + 2026}{4} = 2024,5$$

et

$$y_G = \frac{11,27 + 11,65 + 11,88 + 12,02}{4} = 11,705.$$

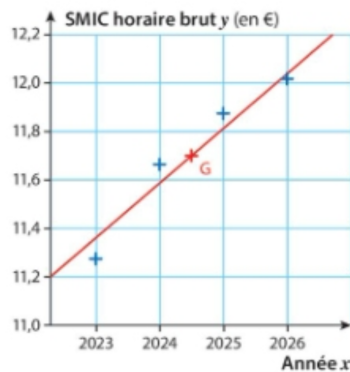


Ajustement affine

Ajuster un nuage de points, c'est trouver une courbe qui passe le plus près possible de tous les points.

Lorsque le nuage de points a une forme suffisamment allongée (points assez bien alignés), on peut l'ajuster par une droite ; on parle alors d'ajustement affine.

Cela permet d'estimer des valeurs intermédiaires (interpolation) ou de faire des prévisions (extrapolation).



On peut tracer une droite d'ajustement « au jugé ». Mais il est possible de tracer des droites différentes.

Il existe des méthodes pour obtenir l'équation réduite d'une droite d'ajustement selon des méthodes précises (méthode de Mayer, méthode des moindres carrés)